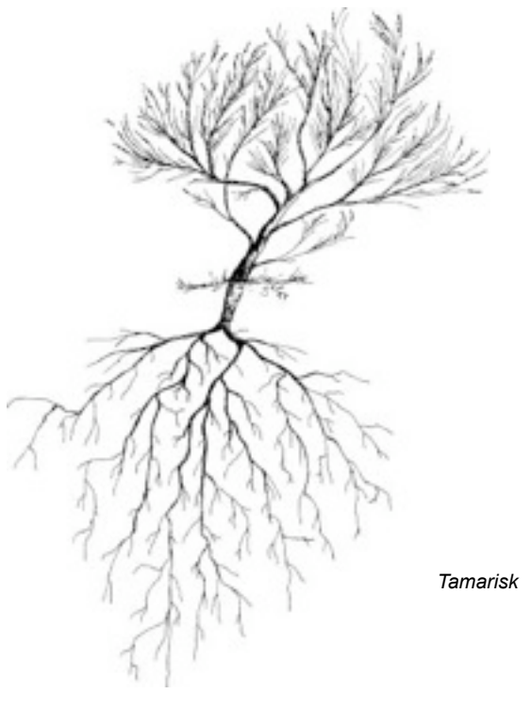
The Invasive Species Forecasting System

A NASA / DOI National Application Project



Tamarisk

John L Schnase

Office of Computational and Information Science and Technology (CISTO / Code 606)

NASA Goddard Space Flight Center

Greenbelt, MD 20771

Outline ...

The ISFS Project

- Phase I - ISFS-in-the-Large
  - USGS Prototype / National Scale
  - National Tamarisk Habitat Suitability Map
- Phase II - ISFS Lite
  - BLM Prototype / Regional Scale
  - Grand Staircase-Escalante National Monument

Implications for Packaging and Distribution ...

- Idiosyncratic diversity
- Regionalization
- Generativity

Invasive Species

An "invasive species" is a species that is non-native to the ecosystem under consideration …

… and whose introduction causes or is likely to cause economic or environmental harm or harm to human health.

- National Invasive Species Council

Invasive Species

A Top Environmental Issue of the 21st Century …

- Economic Costs:
  - *$137+ Billion / Yr*
    (Pimentel, et al. 1999; NISC Management Plan, 2001)
- Environmental Costs:
  - *Decreased biodiversity, ecological services, etc.*
- Human-Health Costs:
  - *West Nile Virus, Malaria, etc.*
- Agricultural Costs:
  - *Crop pathogens, hoof-and-mouth, mad cow disease*

Notorious examples include:

Dutch elm disease, chestnut blight, and purple loosestrife in the northeast; kudzu, Brazilian peppertree, water hyacinth, nutria, and fire ants in the southeast; zebra mussels, leafy spurge, and Asian long-horn beetles in the Midwest; salt cedar, Russian olive, and Africanized bees in the southwest; yellow star thistle, European wild oats, oak wilt disease, Asian clams, and white pine blister rust in California; cheatgrass, various knapweeds and thistles in the Great Basin; whirling disease of salmonids in the northwest; hundreds of invasive species from microbes to mammals in Hawaii; and the brown tree snake in Guam.

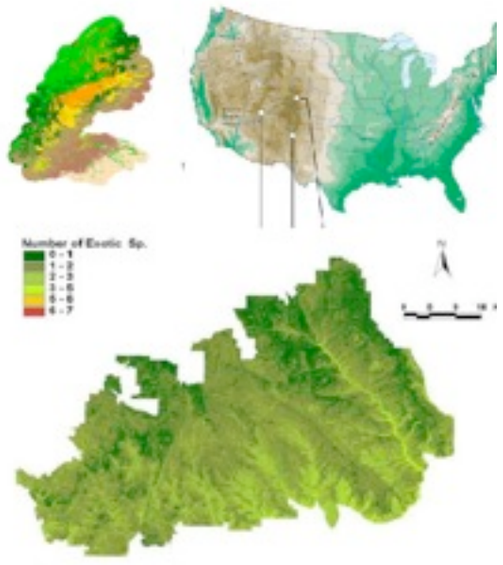
As many as 50,000 now, hundreds new each year …

Federal Government Response

- National Invasive Species Council (EO 13122 - 1999)
- Chaired by USDA, DOI, DOC
- USGS has a lead role in dealing with invasive species science in natural and semi-natural areas

USGS Science / Client Needs

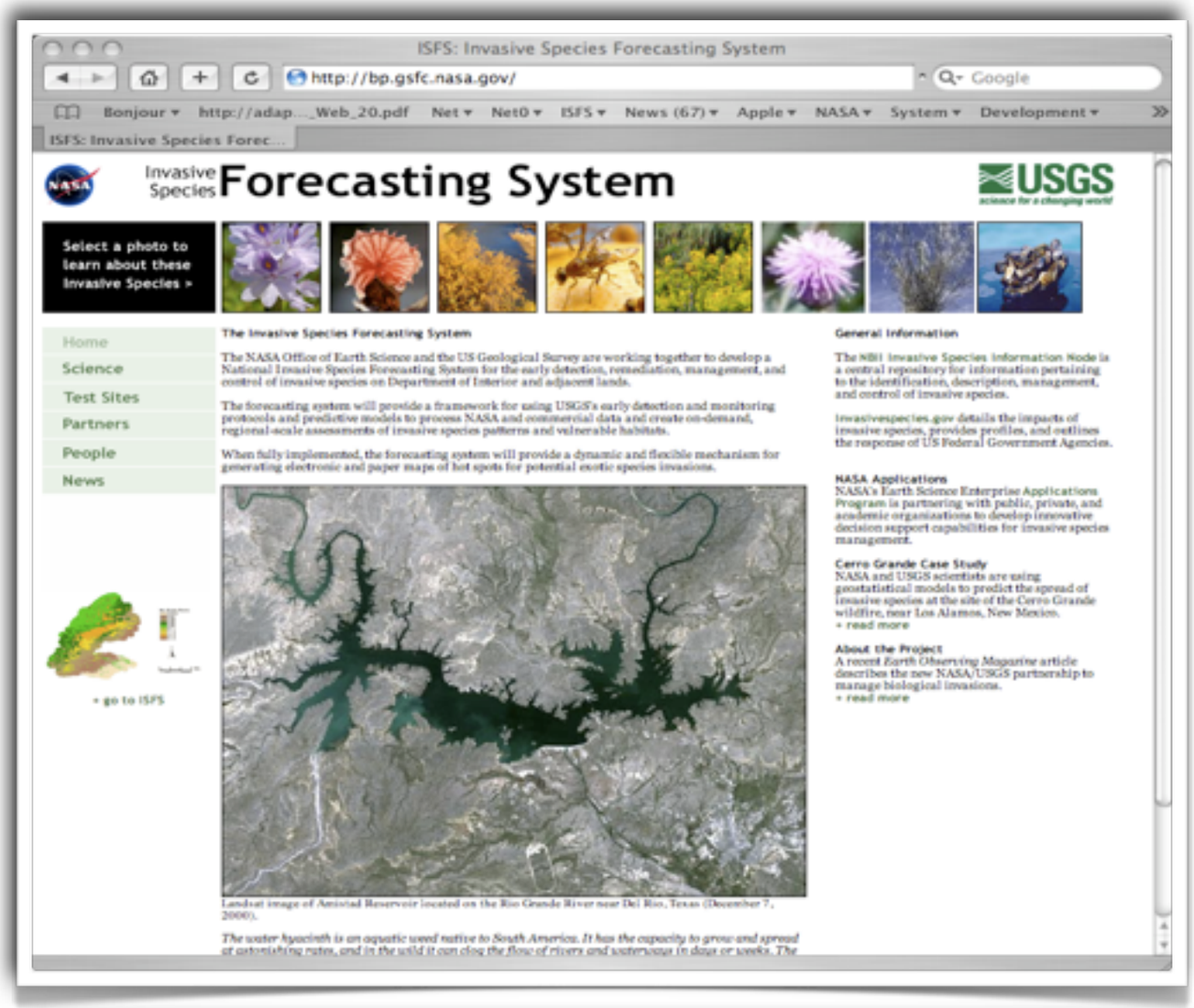
- On-demand, predictive landscape- and regional-scale models and maps for biological invasions
- Low-cost, high-performance computer modeling
- Integrated access to biological field data & NASA Earth Science data



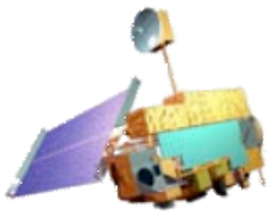
Invasive Species Forecasting System

The ISFS Project ...

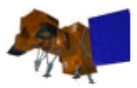
- Partnership between NASA and the Department of Interior
- Goal: Use NASA data and technology to help operational agencies with their invasive species management decision processes.
- Based on USGS's early detection and monitoring protocols.
- Target customers are DOI operational agencies and DOI lands.



Invasive Species Forecasting System



MODIS



LANDSAT

SRTM

Merging continental-scale MODIS time-series data, elevation, soils, and biological field data …

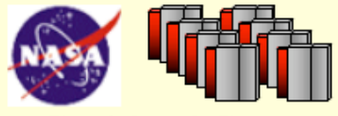
USGS Biological Data

USGS science for a changing world

National Institute for Invasive Species Science

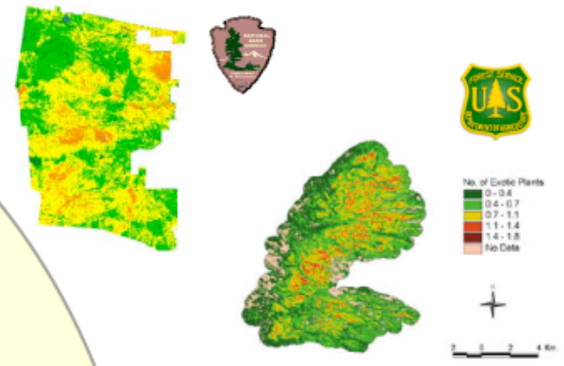
Invasive Species Data Service (ISDS)	Merged Data Set	Invasive Species Analysis & Modeling Service (ISAMS)	Predicted Distribution Map	Invasive Species Decision Support Service (ISDSS)
NBII / ESE Data Access	→	Landscape-Scale Geostatistical Modeling	→	USGS Decision Making

Invasive Species Forecasting System



Goddard Space Flight Center

High-speed networking interconnects NASA/USGS cluster computer, modeling, & mass storage resources …



… to create landscape-scale predictive maps of invasive species distributions.

Invasive Species Forecasting System

Technology Accomplishments

Scalable processing improvements with Cerro Grande Fire Site (CGFS) data

Re-engineered original S-plus code into a Fortran routine ☒

- Reduced processing from 18 days to 61 mins

Parallelized Fortran code ☒

- Reduced processing from 61 mins to 2.47 mins

18 days ☒ 2.5 min changes the science!

Performance of "Adaptive Kriging" for the CGFS study site exceeded goals ☒

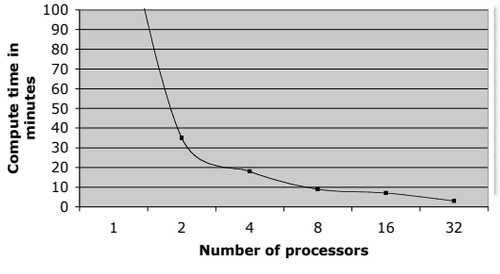
- 1x Area: goal of 2.47 min, achieved 33 seconds. Exceeded goal by 4.5x
- 10x Area: goal is 24.7 min, achieved 4 min 2 sec. Exceeded goal by 6x

Dramatically improved both the quality and capacity of science results for our USGS clients through code optimization and cluster computing

"Constraints in computational time often forced us to substitute simple models for complex, more realistic and accurate models. We needed to greatly reduce computational time to allow us to evaluate larger areas more quickly."

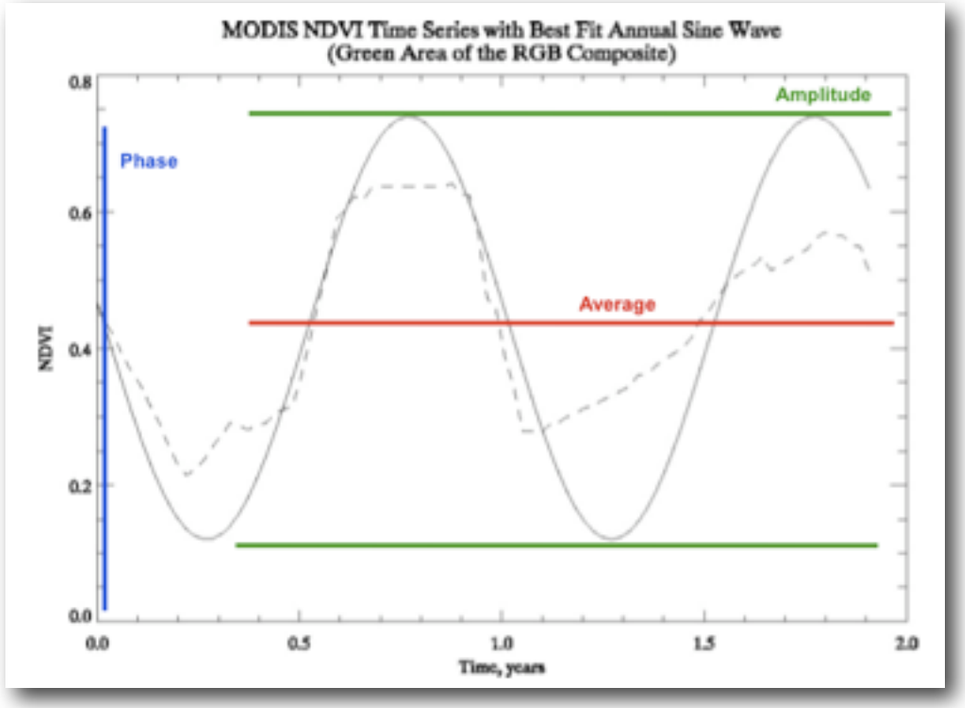
- Dr. Tom Stohlgren, Director National Institute of Invasive Species Science, USGS

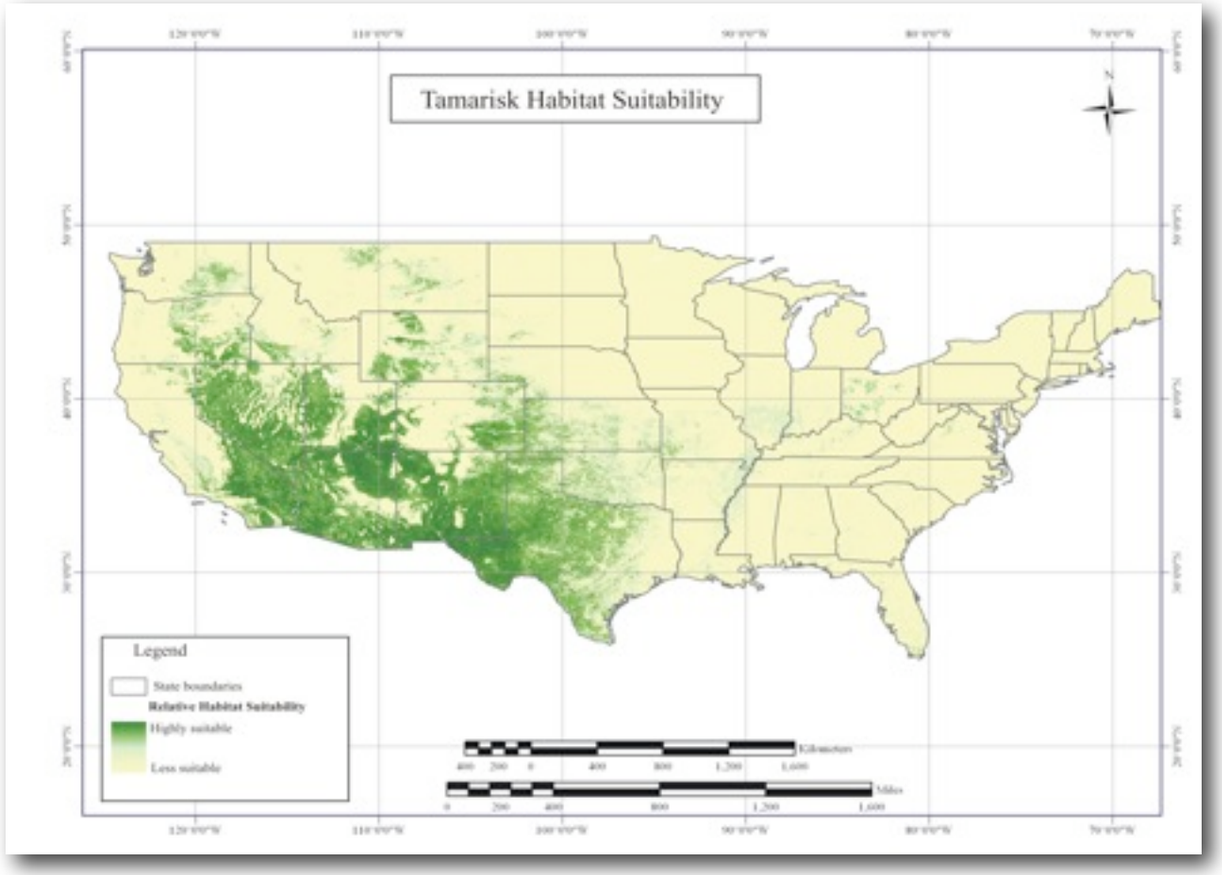
Multiprocessor Scaling Curve
Cerro Grande Fire Site (nn-18)

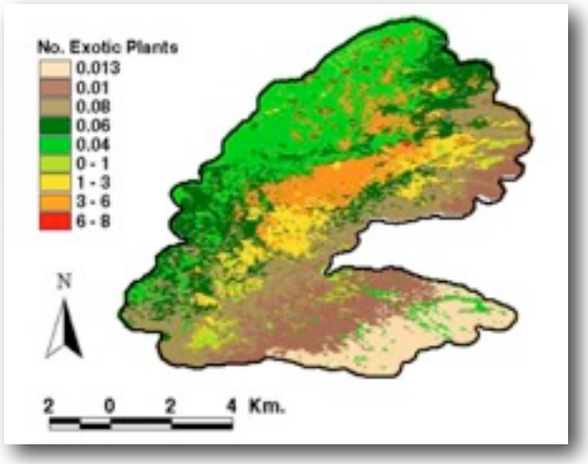


Science Accomplishments ...

- Important ISFS feature is the use of MODIS NDVI time series data
- Seasonal variation in "green up" is an important predictor for many invasive plants
- Time series data used to produce the first national scale habitat suitability map for tamarisk (saltcedar) and in fire ecology studies at the Cerro Grande Wildfire Site near Los Alamos, New Mexico



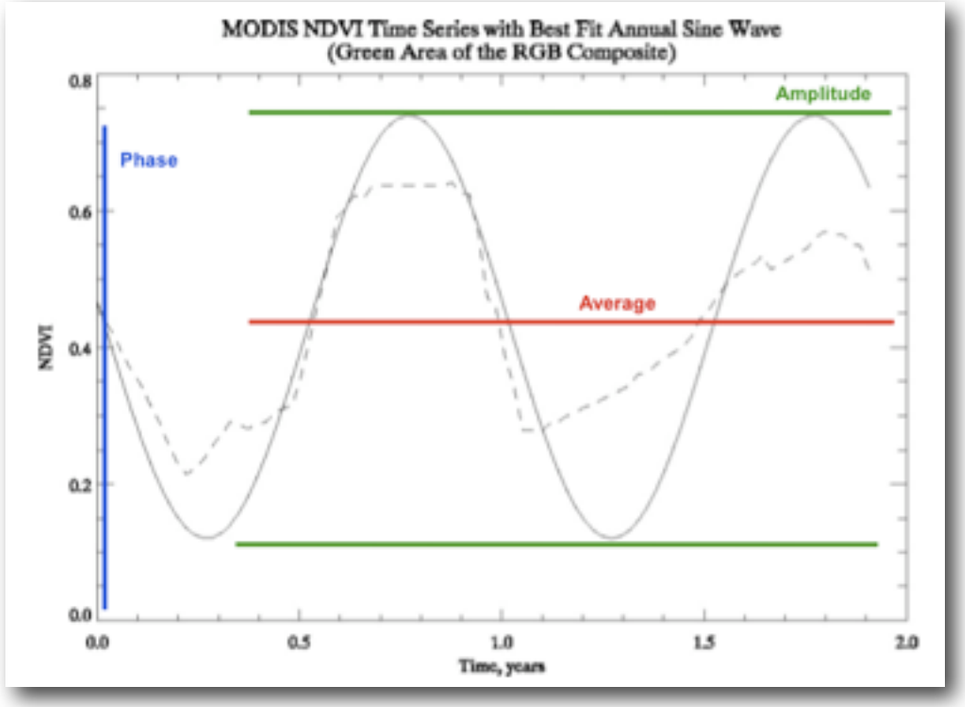


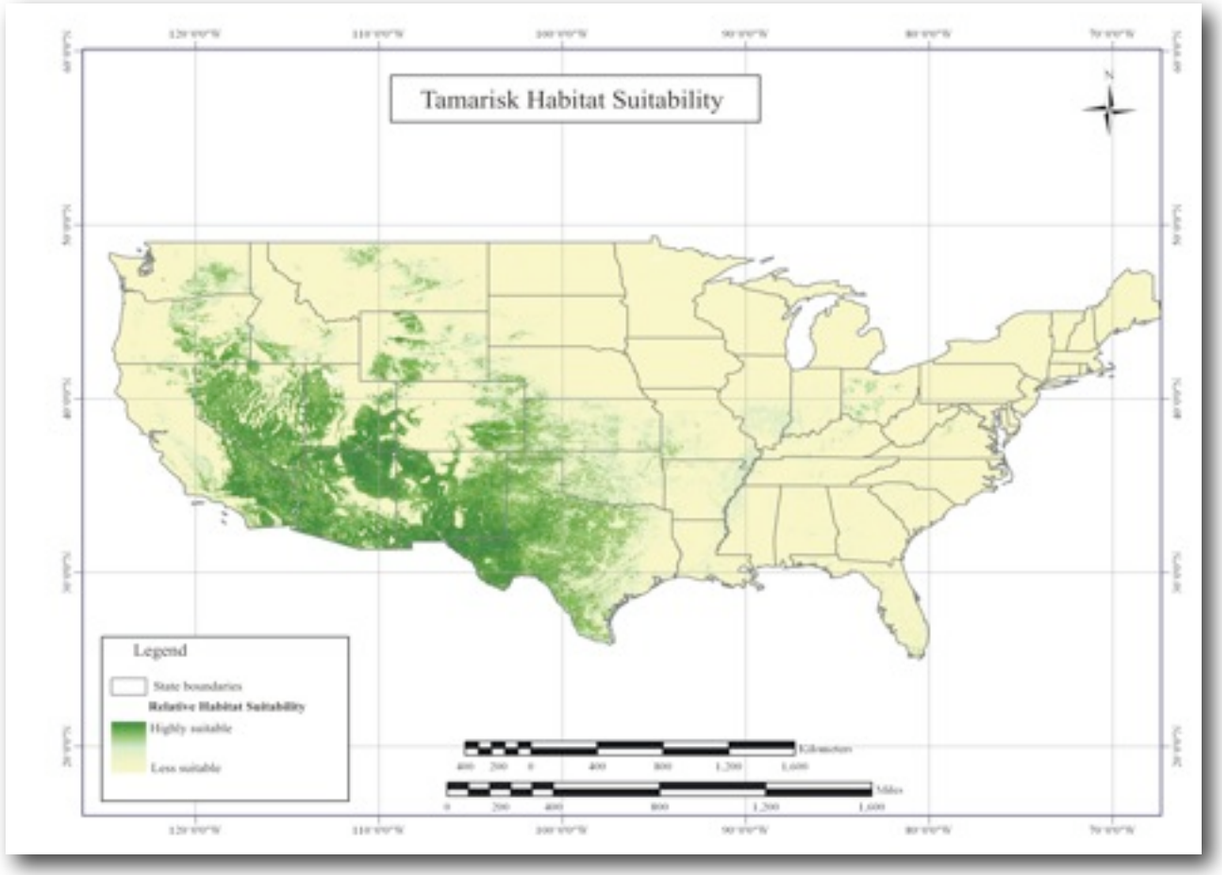


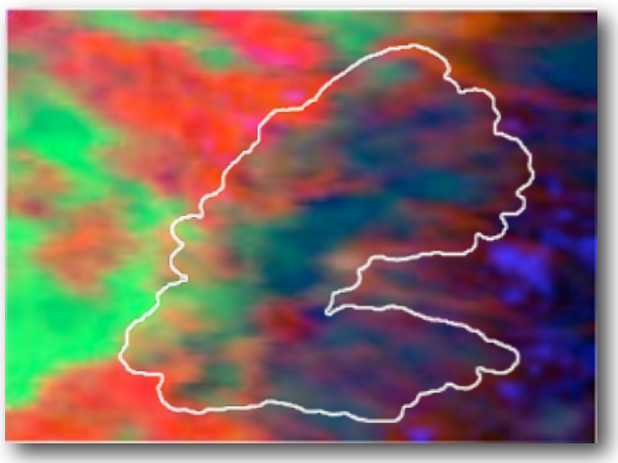
Analysis by Jeff Morisette and Jeff Pedelty, NASA Goddard Space Flight Center

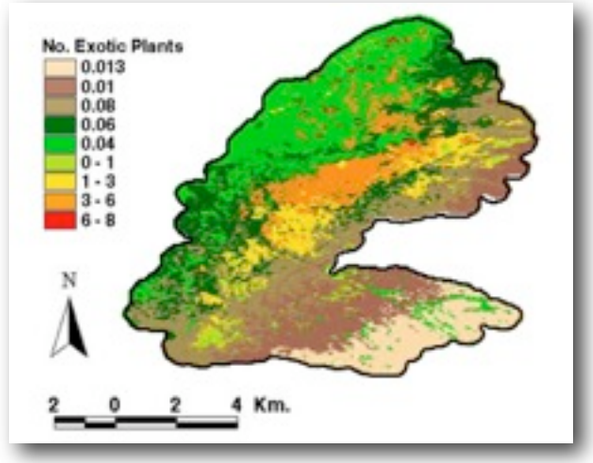
Science Accomplishments ...

- Important ISFS feature is the use of MODIS NDVI time series data
- Seasonal variation in "green up" is an important predictor for many invasive plants
- Time series data used to produce the first national scale habitat suitability map for tamarisk (saltcedar) and in fire ecology studies at the Cerro Grande Wildfire Site near Los Alamos, New Mexico







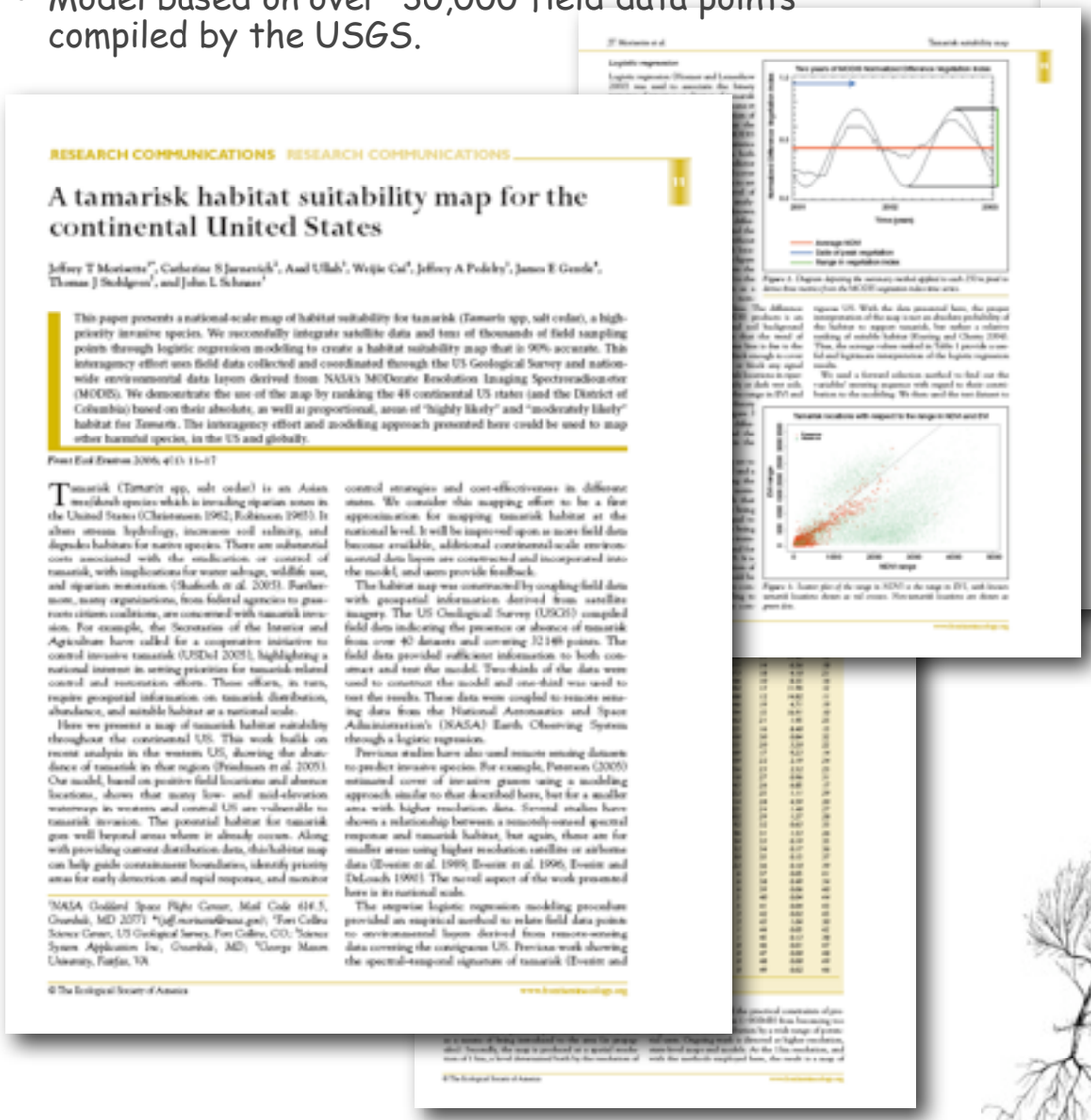


Analysis by Jeff Morisette and Jeff Pedelty, NASA Goddard Space Flight Center


Mapping Invasive Species
Using MODIS Times-Series
Data

Science Accomplishments ...

- National habitat suitability map for tamarisk ...
- A function of MODIS Land Cover and vegetation seasonality.
- Model based on over 30,000 field data points compiled by the USGS.



RESEARCH COMMUNICATIONS RESEARCH COMMUNICATIONS

A tamarisk habitat suitability map for the continental United States

Jeffrey T Morisette[1*], Catherine S Jarnevich[2], Asad Ullah[3], Weijia Cai[4], Jeffrey A Pedelty[5], James E Gentle[6], Thomas J Stohlgren[2], and John L Schnase[1]

This paper presents a national-scale map of habitat suitability for tamarisk (*Tamarix* spp, salt cedar), a high-priority invasive species. We successfully integrate satellite data and tens of thousands of field sampling points through logistic regression modeling to create a habitat suitability map that is 90% accurate. This interagency effort uses field data collected and coordinated through the US Geological Survey and nation-wide environmental data layers derived from NASA's MODerate Resolution Imaging Spectroradiometer (MODIS). We demonstrate the use of the map by ranking the 48 continental US states (and the District of Columbia) based on their absolute, as well as proportional, areas of "highly likely" and "moderately likely" habitat for *Tamarix*. The interagency effort and modeling approach presented here could be used to map other harmful species, in the US and globally.

Front Ecol Environ 2006; 4[1]: 11–17

Tamarisk (*Tamarix* spp, salt cedar) is an Asian woody shrub species which is invading riparian zones in the United States (Christensen 1962; Robinson 1965). It alters stream hydrology, increases soil salinity, and degrades habitats for native species. There are substantial costs associated with the eradication or control of tamarisk, with implications for water salvage, wildlife use, and riparian restoration (Shafroth *et al.* 2005). Furthermore, many organizations, from federal agencies to grassroots citizen coalitions, are concerned with tamarisk invasion. For example, the Secretaries of the Interior and Agriculture have called for a cooperative initiative to control invasive tamarisk (USDoI 2005), highlighting a national interest in setting priorities for tamarisk-related control and restoration efforts. These efforts, in turn, require geospatial information on tamarisk distribution, abundance, and suitable habitat at a national scale.

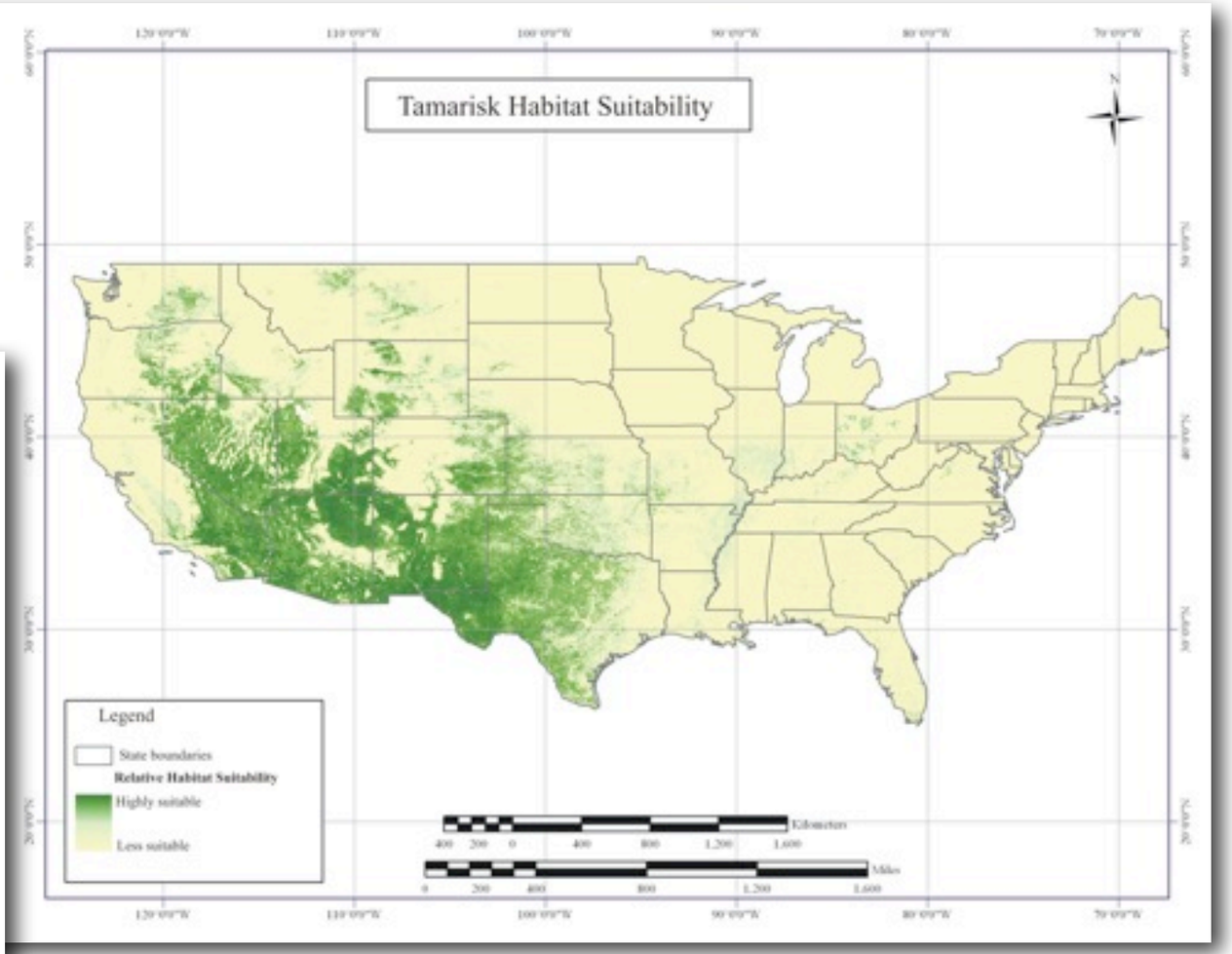
Here we present a map of tamarisk habitat suitability throughout the continental US. This work builds on recent analysis in the western US, showing the abundance of tamarisk in that region (Friedman *et al.* 2005). Our model, based on positive field locations and absence locations, shows that many low- and mid-elevation waterways in western and central US are vulnerable to tamarisk invasion. The potential habitat for tamarisk goes well beyond areas where it already occurs. Along with providing current distribution data, this habitat map can help guide containment boundaries, identify priority areas for early detection and rapid response, and monitor control strategies and cost-effectiveness in different states. We consider this mapping effort to be a first approximation for mapping tamarisk habitat at the national level. It will be improved upon as more field data become available, additional continental-scale environmental data layers are constructed and incorporated into the model, and users provide feedback.

The habitat map was constructed by coupling field data with geospatial information derived from satellite imagery. The US Geological Survey (USGS) compiled field data indicating the presence or absence of tamarisk from over 40 datasets and covering 32 149 points. The field data provided sufficient information to both construct and test the model. Two-thirds of the data were used to construct the model and one-third was used to test the results. These data were coupled to remote sensing data from the National Aeronautics and Space Administration's (NASA) Earth Observing System through a logistic regression.

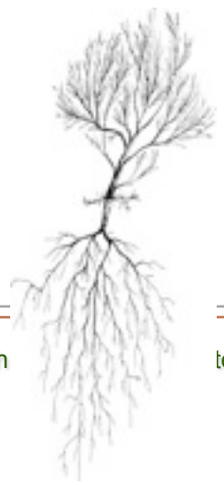
Previous studies have also used remote sensing datasets to predict invasive species. For example, Peterson (2005) estimated cover of invasive grasses using a modeling approach similar to that described here, but for a smaller area with higher resolution data. Several studies have shown a relationship between a remotely-sensed spectral response and tamarisk habitat, but again, these are for smaller areas using higher resolution satellite or airborne data (Everitt *et al.* 1999; Everitt *et al.* 1996; Everitt and DeLoach 1990). The novel aspect of the work presented here is its national scale.

The stepwise logistic regression modeling procedure provided an empirical method to relate field data points to environmental layers derived from remote-sensing data covering the contiguous US. Previous work showing the spectral-temporal signature of tamarisk (Everitt and

[1]NASA Goddard Space Flight Center, Mail Code 614.5, Greenbelt, MD 20771 *(jeff.morisette@nasa.gov); [2]Fort Collins Science Center, US Geological Survey, Fort Collins, CO; [3]Science System Application Inc, Greenbelt, MD; [6]George Mason University, Fairfax, VA



Morisette, J.T., C. S. Jernevich, A. Ullah, W. Cai, J.A. Pedelty, J. Gentle, T.J.Stohlgren, J.L. Schnase, A tamarisk habitat suitability map for the continental US., *Frontiers in Ecology*, February 2006.



Phase I

"ISFS-in-the-Large"

A classic implementation:
A physically, programmatically, and
ideologically centralized Web service ...

Phase I

"ISFS-in-the-Large"

But what happens
when the context changes?

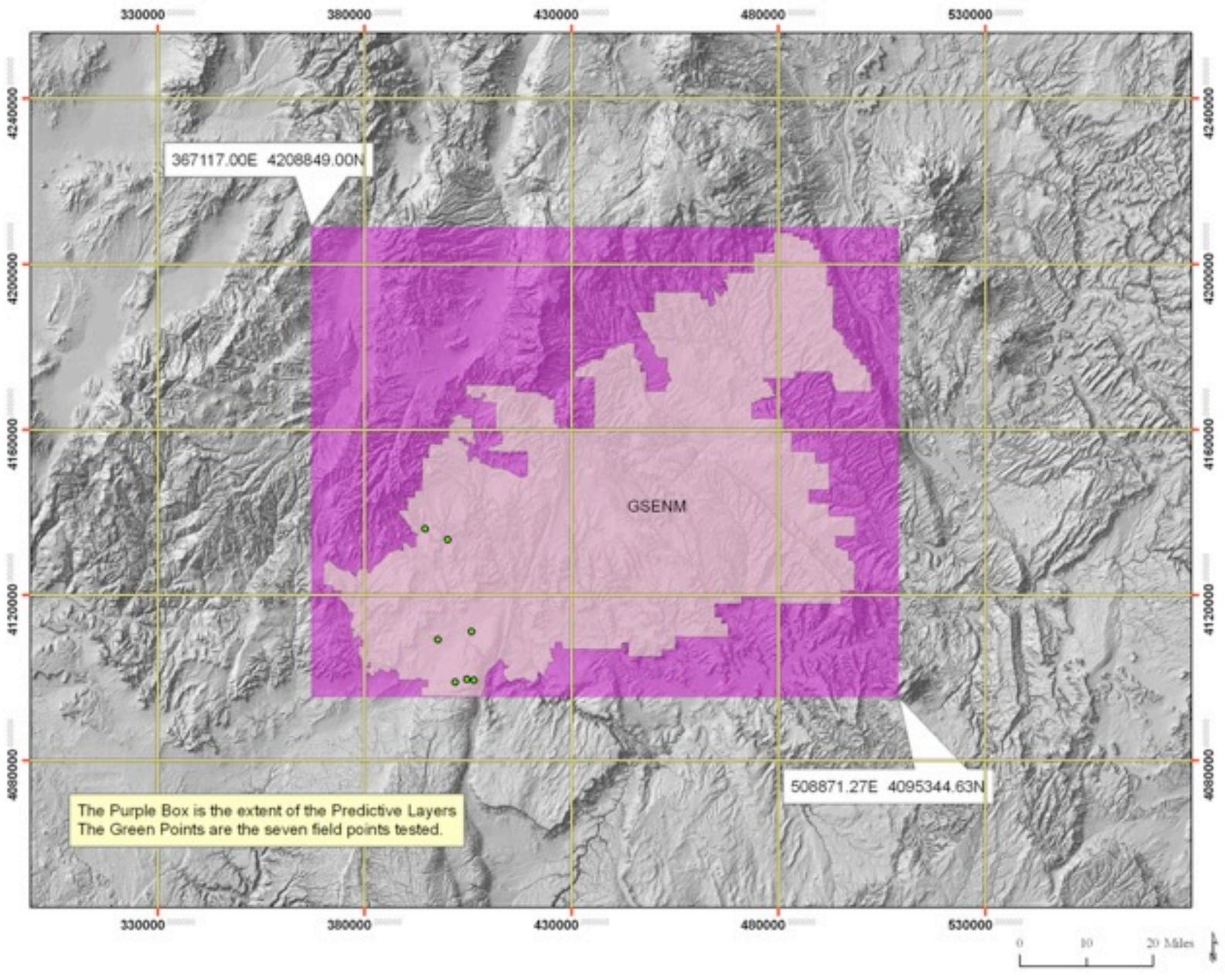
Phase II

"ISFS Lite"

What's the simplest possible
way of doing something useful?

What is the simplest possible way of doing something useful?

1. **Limit the application to a region of interest**
2. Input p/a field data for species of interest
3. Input site-specific environmental predictors
4. Produce predicted habitat suitability map

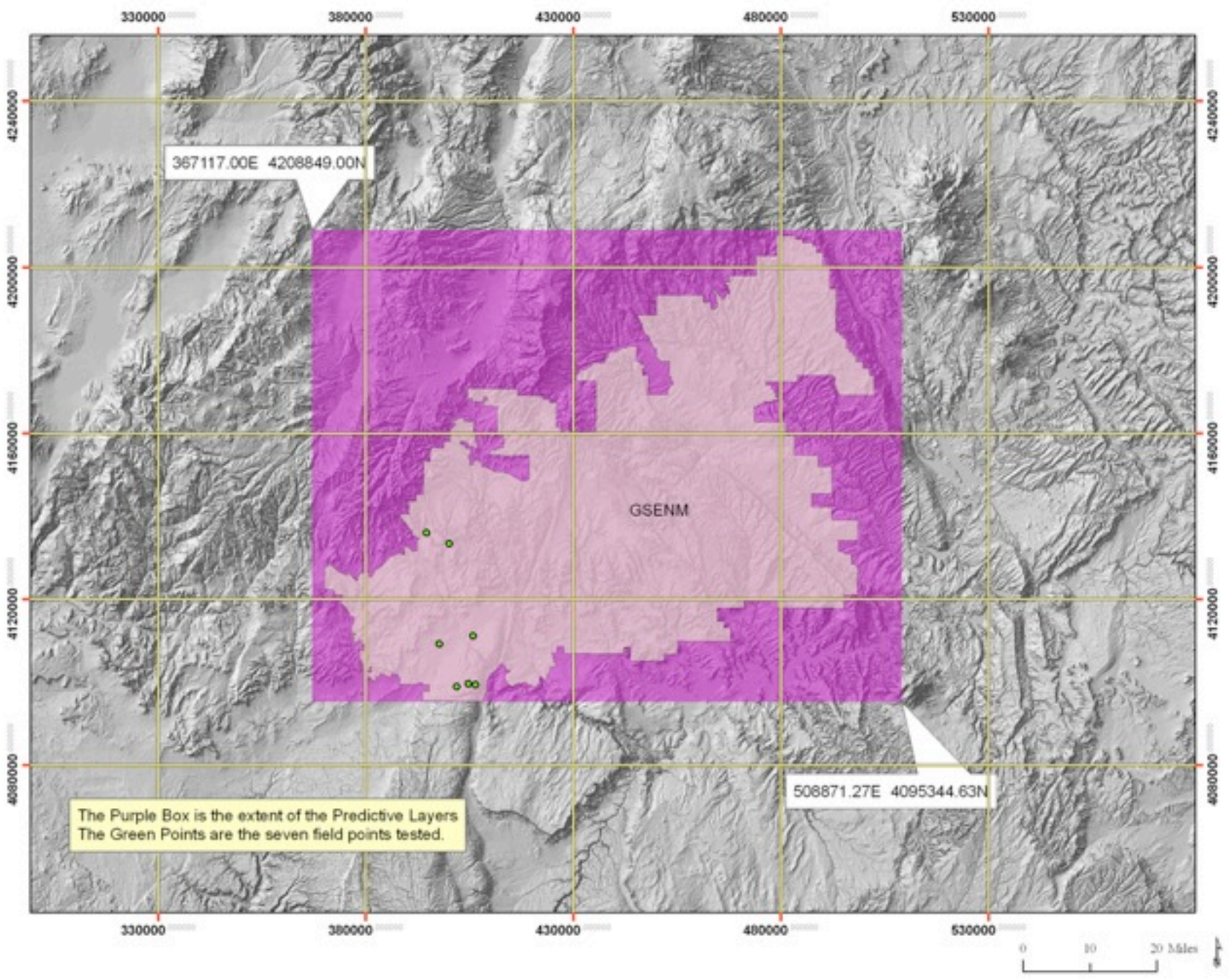


xutm	yutm	pres_abs
422424	4159674	1
405105	4127066	1
404229	4127179	1
434927	4174629	0
423256	4161403	0
459478	4111353	1
430958	4112763	1
417839	4105218	1
421461	4109301	1
442243	4103983	1
436867	4141688	0
432029	4148220	0
433270	4147739	0
[illegible]875	4141782	0
[illegible]7	4131352	0
	4130866	1
	4138089	0
	4137960	0
	4133851	1
	4130477	0
	4102182	1

CSV

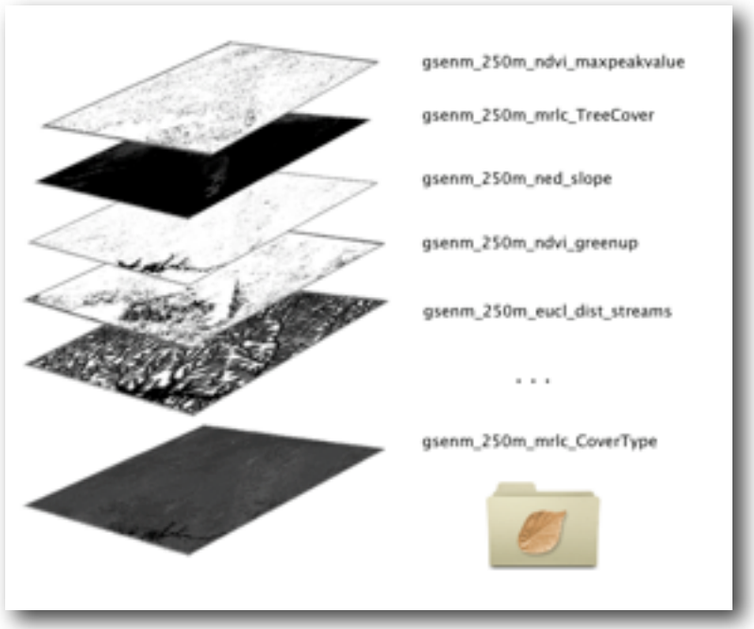
What is the simplest possible way of doing something useful?

1. Limit the application to a region of interest
2. **Input p/a field data for species of interest**
3. Input site-specific environmental predictors
4. Produce predicted habitat suitability map



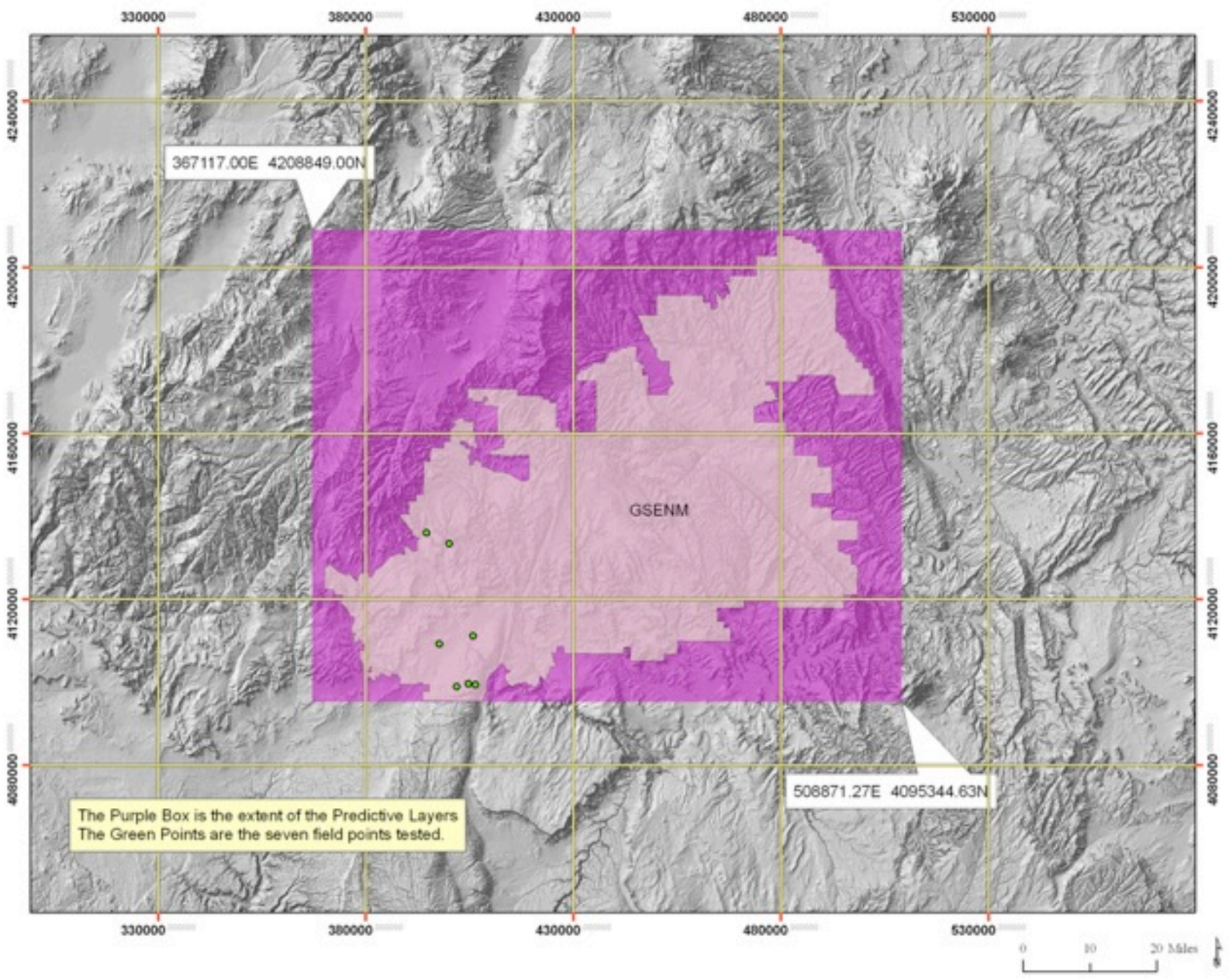
xutm	yutm	pres_abs
422424	4159674	1
405105	4127066	1
404229	4127179	1
434927	4174629	0
423256	4161403	0
459478	4111353	1
430958	4112763	1
417839	4105218	1
421461	4109301	1
442243	4103983	1
436867	4141688	0
432029	4148220	0
433270	4147739	0
[illegible]875	4141782	0
[illegible]7	4131352	0
	4130866	1
	4138089	0
	4137960	0
	4133851	1
	4130477	0
	4102182	1

CSV



What is the simplest possible way of doing something useful?

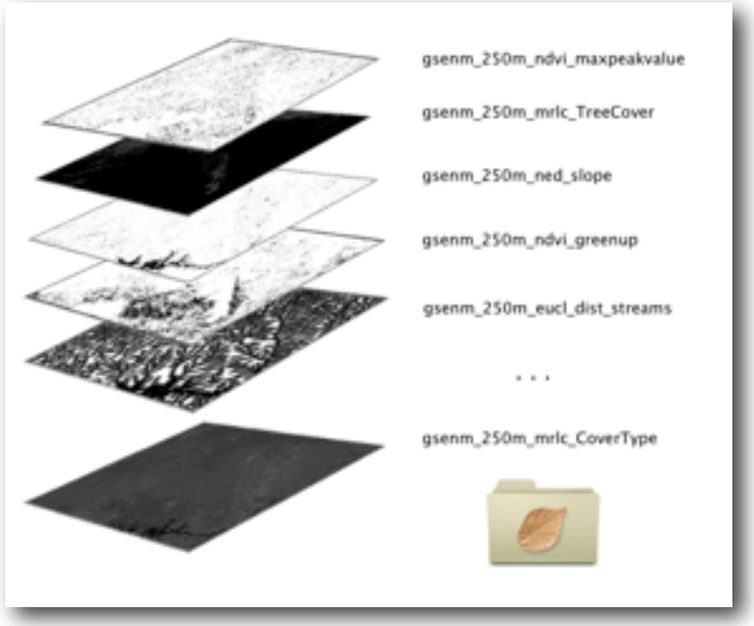
1. Limit the application to a region of interest
2. Input p/a field data for species of interest
3. **Input site-specific environmental predictors**
4. Produce predicted habitat suitability map



Invasive Species Forecasting System

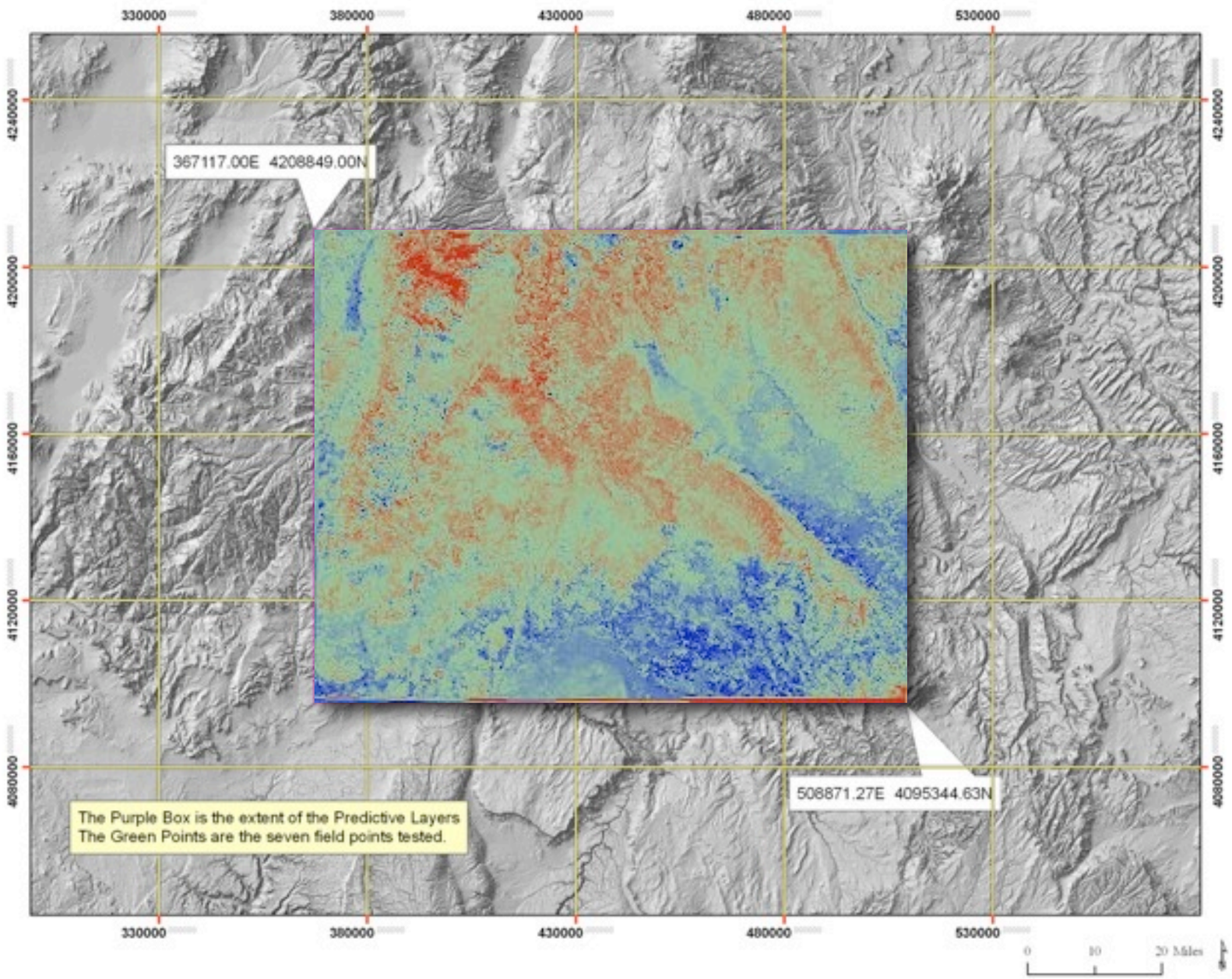
xutm	yutm	pres_abs
422424	4159674	1
405105	4127066	1
404229	4127179	1
434927	4174629	0
423256	4161403	0
459478	4111353	1
430958	4112763	1
417839	4105218	1
421461	4109301	1
442243	4103983	1
436867	4141688	0
432029	4148220	0
433270	4147739	0
[illegible]875	4141782	0
[illegible]7	4131352	0
	4130866	1
	4138089	0
	4137960	0
	4133851	1
	4130477	0
	4102182	1

CSV



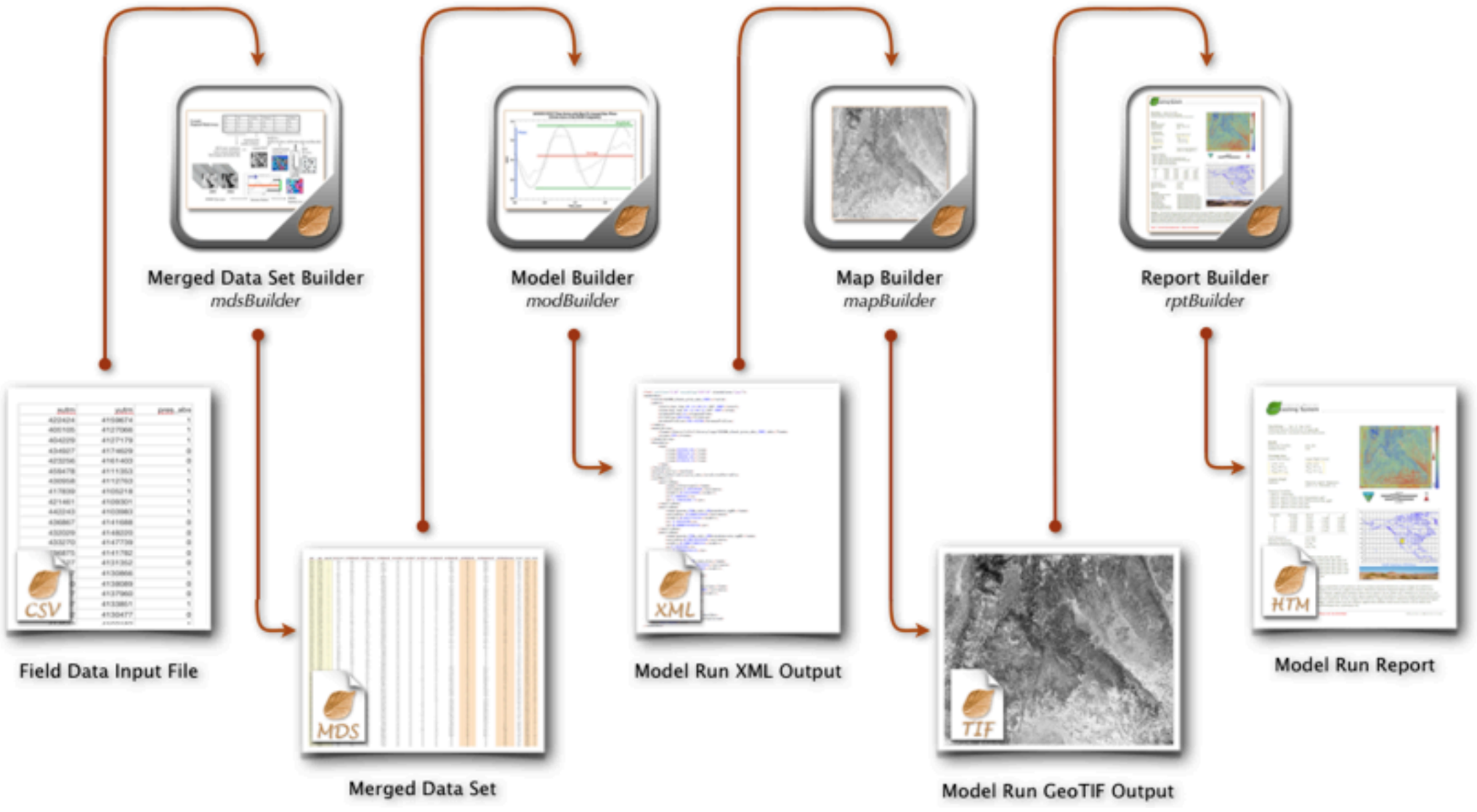
What is the simplest possible way of doing something useful?

1. Limit application to a region of interest
2. Input p/a field data for species of interest
3. Input site-specific environmental predictors
4. **Produce predicted habitat suitability map**



What is the simplest possible way of doing something useful?

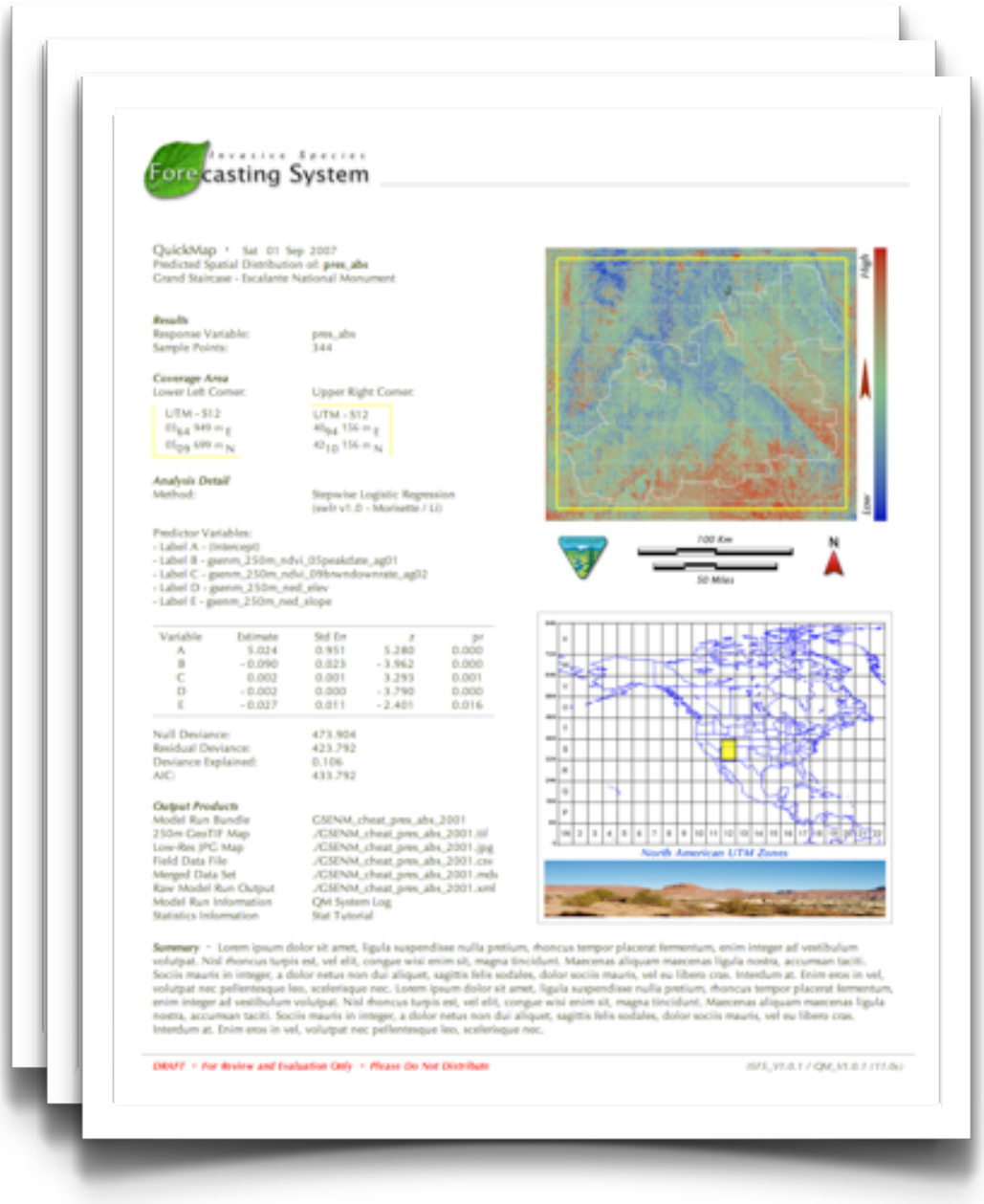
2. Input p/a field data for species of interest
3. Input site-specific environmental predictors
4. Produce predicted habitat suitability map
5. **Use simple, hardened workflow components**



Invasive Species Forecasting System

What is the simplest possible way of doing something useful?

3. Input site-specific environmental predictors
4. Produce predicted habitat suitability map
5. Use simple, hardened workflow components
6. **Use minimal interfaces and minimal outputs**



Invasive Species Forecasting System

What is the simplest possible way of doing something useful?

4. Produce predicted habitat suitability map
5. Use simple, hardened workflow components
6. Use minimal interfaces and minimal outputs
7. **Accommodate multiple platforms and uses**

Invasive Species Forecasting System

ISFS / QuickMap @ Grand Staircase-Escalante National Monument

A light-weight, site-specific runtime service ...

Based on a simple, adaptable, componentized "ISFS Framework" that implements a generative approach to scientific work-flow management ...

Light-weight application, runs on laptop and desktop computers, assumes only intermittent / asynchronous internet connectivity, personalized / private data management, iTunes U / RSS / peer-to-peer sharing and communication ...

ISFS-G100-Console
ISFS-G101-Moqui
ISFS-G102-Cayote
ISFS-G103-Peregrine
ISFS-G104-Pareah
ISFS-G105-Wolverine

Invasive Species Forecasting System

LISN / NPS Fire Ecology / USGS Development

Invasive Species Modeling and Assessment System

Global Organism Detection and Monitoring System

Naitonal Biological Information Infrastructure

USGS Fort Collins Science Center

ISFS Development

Current Situation

- Two complementary threads
- One shared prime contractor (Located in JLS lab at GSFC)
- Same people, different places (Shasby, Morisette, Sheffner, etc.)
- Synergistically intertwingled YERC, NPS, USGS, NASA GSFC, NASA ARC activities ...

2009

Julian

2008

ISFS Framework (V1.65)

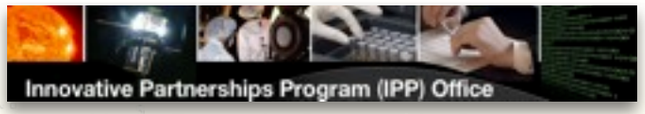
2006

ISFS Framework (V1.0-V1.5)

2005

2000

ISFS J2EE Engineering Prototype

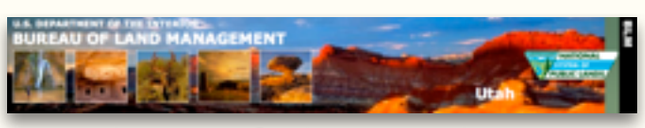


NASA Innovative Partnerships Program

Burned Area Emergency Response

National Lanscape Conservation System

National Geospatial Technology Extension Network



Grand Staircase-Escalante National Monument

USGS -
Web-based analysis services ...

- ISFS component technology and data products being integrated into several online modeling and data services
- Online ISFS modeling capabilities available to the National Park Service
- Coordination thru Innovim, LLC, the prime contractor for NASA and USGS
- Jeff Morisette (former NASA collaborator) now at USGS facilitating invasive species research and modeling ...

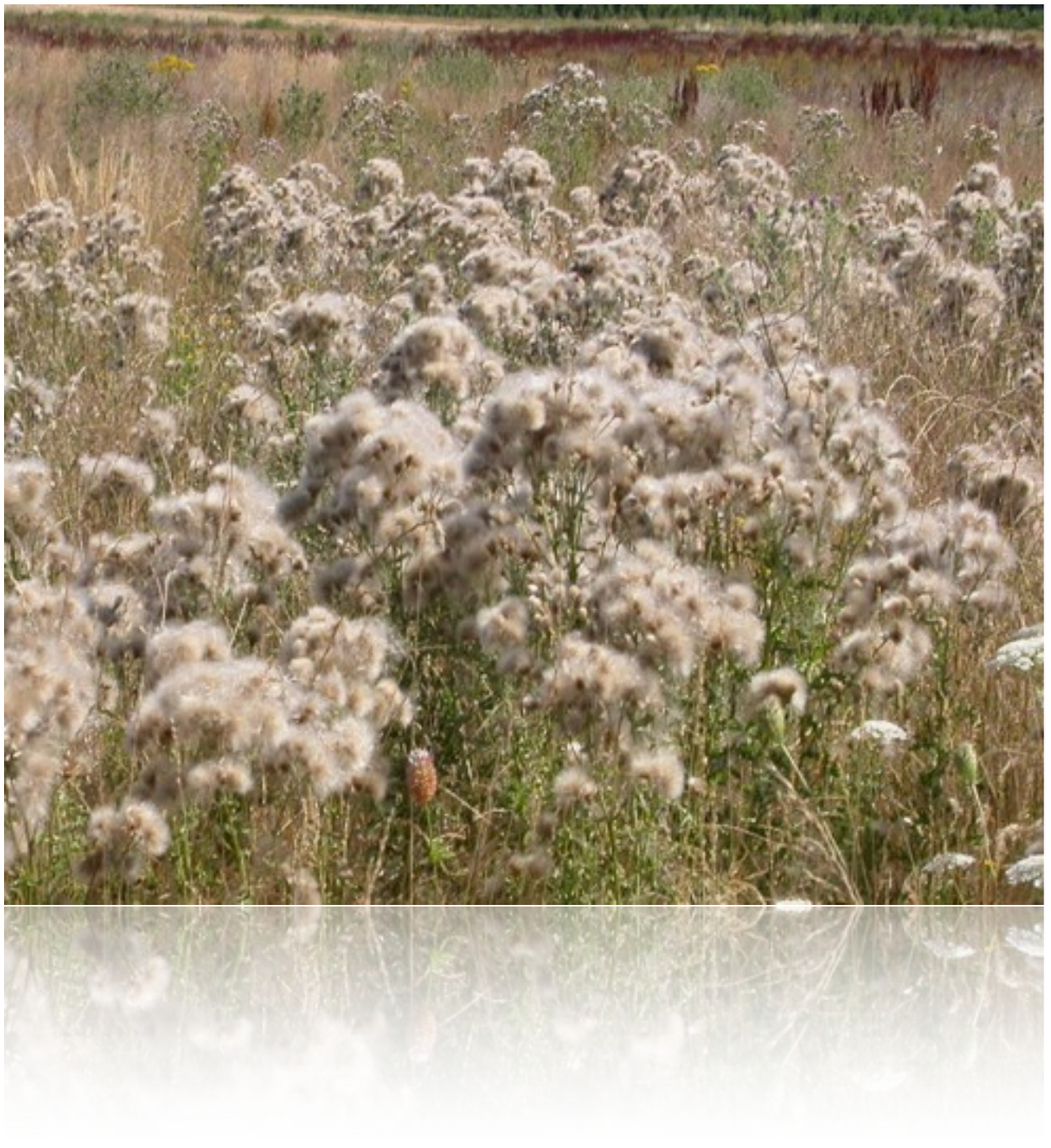
NASA -
Web-enabled DSS framework and tools ...

- Operational deployment at BLM's Grand Staircase-Escalante National Monument
- New collaborators include NASA ARC, BLM's National Landscape Conservation System, and the National Geospatial Technology Extension Network
- Being evaluated for use by NIFC's Burned Area Emergency Response Program
- Subject of 7 NASA New Technology Reports
- Technology available for government and private-sector licensing and partnering

So what?

Implications for the packaging and distribution of assets ...

- ISFS representative of many decision support problems in biodiversity, ecosystem forecasting, natural resource management
- NASA does science and data at a global scale, but the vast majority of policy, budgetary, and resource management decision-making occurs at a regional scale
- For NAS's "societal benefit" goals of the Decadal Survey missions, think about applications that enable the social construction of scientific practice



So what?

Implications for the packaging and distribution of assets ...

- ISFS representative of many decision support problems in biodiversity, ecosystem forecasting, natural resource management
- NASA does science and data at a global scale, but the vast majority of policy, budgetary, and resource management decision-making occurs at a regional scale
- For NAS's "societal benefit" goals of the Decadal Survey missions, think about applications that enable the social construction of scientific practice

Unique, but ...

- applicable to a large customer base in this domain
- applicable to many related biological problems
- implements frequently encountered workflow processes of data assembly, data integration, analysis, presentation rendering, publishing, collaboration
- involves interdisciplinary, inter- and intra-organizational partnerships

... idiosyncratic diversity

So what?

Implications for the packaging and distribution of assets ...

- ISFS representative of many decision support problems in biodiversity, ecosystem forecasting, natural resource management
- NASA does science and data at a global scale, but most policy, budgetary, and resource management decision-making occurs at a regional scale
- For NAS's "societal benefit" goals of the Decadal Survey missions, think about applications that enable the social construction of scientific practice

Perhaps the single most significant shift ...

- "societal benefit" means greater direct societal involvement
- direct societal involvement means a hugely expanding customer base
- iTunes : Playlist :: Decadal Survey Data : Emerging DS Applications
- implies increased localization, specialization, personalization

... regionalization, tailorability

So what?

Implications for the packaging and distribution of assets ...

- ISFS representative of many decision support problems in biodiversity, ecosystem forecasting, natural resource management
- NASA does science and data at a global scale, but the vast majority of policy, budgetary, and resource management decision-making occurs at a regional scale
- For NAS's "societal benefit" goals of the Decadal Survey missions, think about applications that enable the social construction of scientific practice

Generativity refers to a system's capacity to produce unanticipated change through unfiltered contributions from broad and varied audiences. The concept highlights aspects of an innovation or process that enable an autocatalytic feeding-forward — a self-assembly — that can help make growth, further innovation, and success possible.

Zittrain identifies five properties of generative systems:

(1) How extensively a system or technology leverages a set of possible tasks: **Leverage** makes a difficult job easier, and, in general, the more a system can do, the more capable it is of producing change.
(2) How well it can be adapted to a range of tasks: **Adaptability** enables new, unintended, and innovative uses of a technology. It broadens the technology's use.
(3) How easily new contributors can master it: **Ease of Mastery** reflects how easy it is for broad audiences to understand how to adopt and adapt it. The more useful a technology is both to the neophyte and the expert, the more generative it is.
(4) How accessible it is to those ready and able to build on it: **Accessibility** makes it easier to obtain the technology and the information necessary to achieve mastery. The more accessible, the more generative.
(5) How transferable any changes are to others, including non-experts: **Transferability** reflects how easily changes in the technology can be conveyed to others.

... generativity

Zittrain, J. 2008. *The Future of the Internet—And How to Stop It*. Yale University Press, New Haven. 342 pp. This material is also available on the Web. *See* the Future of the Internet Posting, http://yupnet.org/zittrain/archives/13#11 (as of August 21, 2008, 10:30 EST).

Invasive Species
Forecasting System

GRAND STAIRCASE - ESCALANTE
NATIONAL MONUMENT